

OBSERVATIONS FOR REANALYSES

STEFAN BRÖNNIMANN, ROB ALLAN, CHRISTOPHER ATKINSON, ROBERTO BUIZZA, OLGA BULYGINA, PER DAHLGREN, DICK DEE, ROBERT DUNN, PEDRO GOMES, VIJU O. JOHN, SYLVIE JOURDAIN, LEOPOLD HAIMBERGER, HANS HERSBACH, JOHN KENNEDY, PAUL POLI, JOUNI PULLIAINEN, NICK RAYNER, ROGER SAUNDERS, JÖRG SCHULZ, ALEXANDER STERIN, ALEXANDER STICKLER, HOLLY TITCHNER, MARIA ANTONIA VALENTE, CLARA VENTURA, AND CLIVE WILKINSON

Reanalysis efforts depend on observations, and the corresponding work feeds into public repositories, ideally occurring one cycle ahead of reanalysis generation.

Reanalysis efforts fundamentally depend on observations of the atmosphere, the ocean, the cryosphere, and the land surface. First and foremost, observations provide information on the state of the atmosphere and ocean that can be assimilated into a numerical weather prediction model in order to produce the reanalysis. However, observations are also used in several other steps along the processing chain. They are used to constrain the boundary

conditions of the numerical weather prediction model, to calibrate statistical relations used in the processing (e.g., for geophysical parameter estimation from satellite data), to determine and correct the error of other observations, and to evaluate the final reanalysis product. A reanalysis is therefore only as good as the underlying observational data, and rescuing, compiling, and managing additional observations are therefore critical.

AFFILIATIONS: BRÖNNIMANN AND STICKLER—Oeschger Centre for Climate Change Research, and Institute of Geography, University of Bern, Bern, Switzerland; ALLAN, ATKINSON, DUNN, KENNEDY, RAYNER, SAUNDERS, AND TITCHNER—Met Office, Exeter, United Kingdom; BUIZZA, DAHLGREN, DEE, HERSBACH, AND POLI—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom; BULYGINA AND STERIN—Russian Research Institute of Hydrometeorological Information, Obninsk, Russia; GOMES, VALENTE, AND VENTURA—Instituto Dom Luiz Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal; JOHN AND SCHULZ—European Organisation for the Exploitation of Meteorological Satellites, Darmstadt, Germany; JOURDAIN—Météo-France, Paris, France; HAIMBERGER—University

of Vienna, Vienna, Austria; PULLIAINEN—Finnish Meteorological Institute, Helsinki, Finland; WILKINSON—University of East Anglia, Norwich, United Kingdom

CORRESPONDING AUTHOR: Stefan Brönnimann, stefan.broennimann@giub.unibe.ch

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-17-0229.1

In final form 13 February 2018
©2018 American Meteorological Society



This article is licensed under a [Creative Commons Attribution 4.0 license](#).

Collecting historical observations is a tremendous effort, which can only be undertaken through close collaboration between weather and climate data centers, satellite agencies, reanalysis producers, and individual scientists. Reanalysis producers share their observation databases and thus build on each other's efforts. However, each reanalysis project also has its specific objectives and technical restrictions or opportunities. Particularly when going back in time, not all observations are easily accessible. A large fraction of historical meteorological observations has never been digitized because the data have either thus far not been considered valuable or just not been known about. However, even in the rather recent past, the availability of satellite data records (and the computer codes needed to read and process them) is an issue that needs to be addressed (Poli et al. 2017). Furthermore, metadata need to be compiled and shared in a systematic way in order to render the data useful.

Major efforts were therefore undertaken in the European Reanalysis of the Global Climate System (ERA-CLIM2), to collect and make available observations for all reanalyses covering the period since 1900. Such an undertaking requires a much broader vision than the production of a specific reanalysis. Historical observations are also a legacy, and generating reanalyses (or other data products) must be seen as a continuous effort. The observations collected may be fully used only in the next cycle of reanalysis generation. However, by simulating the observations already in the current generation of reanalyses, feedback on their quality can be obtained and many instrument issues have been detected this way (Lawrence et al. 2017).

In this paper we report on the efforts within ERA-CLIM2 to rescue, compile, assess, and make available observations as well as metadata and expertise. The data generated feed into global repositories and serve future reanalysis projects.

CLIMATE REANALYSES. Reanalyses are generated by combining first-guess estimates of the state of the atmosphere, ocean cryosphere, land, and so forth, usually defined by model forecasts, with data from a range of observing platforms (surface, upper air, satellites) using objective methods named data assimilation systems. The availability and quality of observations affects the confidence that can be placed upon reanalysis datasets (e.g., Compo et al. 2011; Hersbach et al. 2017; Uppala et al. 2005).

Some reanalyses provide multidecadal or longer time series of gridded variables. However, they can

only be used to monitor the evolution of the Earth system climate if they are not sensitive to changes of the observing system. Given that the observing system has been changing dramatically over the past decades, especially since the 1970s with the start of the satellite era, it is necessary to try to rescue and recover as many past observations as possible, so that the number of observations does not change too strongly throughout the years spanned by a climate reanalysis and so that different observations can be used together to characterize their biases or other uncertainties.

The work presented here was performed within the framework of ERA-CLIM2, a 4-yr research project funded by the European Union Seventh Framework Program (FP7). It aims to produce coupled reanalyses, which are physically consistent datasets describing the evolution of the global atmosphere, ocean, land surface, cryosphere, and carbon cycle [see Buizza et al. (2018) for an overview]. The main contribution of the ERA-CLIM2 project to climate science has been to improve the capacity for producing state-of-the-art climate reanalyses that extend back to the early twentieth century. Observational data rescue was an important part of the project, including data rescue of historic in situ weather observations around the world and reprocessing of satellite climate data records. Specifically, the work encompassed 1) data rescue for in situ observations, their quality control, and provision of metadata; 2) satellite data rescue, reprocessing, and intercalibration; and 3) provision of boundary constraints and external forcing.

SURFACE METEOROLOGICAL DATA.

Observations of pressure and wind and, sometimes, temperature at the surface are assimilated into reanalyses. Other surface observations, such as those of precipitation, have so far mostly been used in reanalysis to evaluate the effectiveness of the forecast of those variables. It is important to apply rigorous quality control (QC) to these observations and the Hadley Centre Integrated Surface Database (HadISD; version 2.0.0; Dunn et al. 2016; available from www.metoffice.gov.uk/hadobs/hadisd/) has been further developed under ERA-CLIM2 to improve and extend the QC of a subset of subdaily observations back to 1931 obtained from the National Oceanic and Atmospheric Administration/National Centers for Environmental Information (NOAA/NCEI) Integrated Surface Database (Smith et al. 2011), as well as providing an assessment of station record homogeneity.

However, existing holdings of such observations are inadequate, and ERA-CLIM2 contributed to the

collection of further historical surface meteorological data by cataloging, prioritizing, imaging, digitizing, applying quality-control tools, and/or formatting subdaily data from various sources (see Table 1). In addition to data rescue work carried out within the project, ERA-CLIM2 supported the coordination of global data rescue work via the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative (Allan et al. 2011). In total, 4.7 million station days' worth of meteorological data (this does not include the snow data discussed later) were cataloged. Most of the data (4.6 million station days) were digitized (Table 1). Work was also performed on the QC and homogenization of the data (e.g., Bližňák et al. 2015; Hunziker et al. 2017). The data were submitted to the corresponding data centers, such as the International Surface Pressure Databank (ISPD; Cram et al. 2015) for pressure and the International Surface Temperature Initiative (ISTI; Rennie et al. 2014) for temperature; additional download links for individual data collections are given below. Additionally, metadata for these stations are included in the ERA-CLIM registry (see below). Table 1 provides a brief listing of important sources. These are discussed in more detail below.

Colónia de Angola—Elementos Meteorológicos e Climatológicos, 1937–74. These publications presented data for Luanda from 1937 until 1974 (with a gap in 1969) and for other Angolan stations from 1953. The entire collection was imaged and subsequently recovered by optical character recognition (OCR) techniques; ERA-CLIM2 complemented the work started in ERA-CLIM (Stickler et al. 2014). The subdaily data chosen for digitization included surface or mean sea level (MSL) pressure, air temperatures, relative humidity, wind speed and direction, cloud cover, precipitation, and evaporation. In addition to Luanda, data from nine stations in Angola were digitized for ERA-CLIM2: Mossâmedes (now Namibe), Cabinda, Dundo, Malange, Vila Lusa (now Luena), Lobito, Nova Lisboa (now Huambo), Sá da Bandeira (now Lubango), and Mavinga. Subsequent to the OCR process, QC tools were applied to the digitized data. The data can be accessed online (<http://eraclim2.rd.ciencias.ulisboa.pt/>).

TABLE 1. Surface observations (No. of station days) digitized within ERA-CLIM and ERA-CLIM2.

| Source | Cataloged | Digitized | Quality controlled |
|---|-----------|-----------|--------------------|
| Backward extension (<1965) of meteorological data from 246 Russian stations | 2,738,595 | 2,738,595 | 2,738,595 |
| 41 Chilean stations, 1950–99 | 383,151 | 357,456 | 36,682 |
| 76 Portuguese stations in Portugal and ex-colonies in Africa and Asia | 1,020,727 | 1,009,131 | 605,478 |
| South China Sea logbooks for 100 stations | 830,286 | 830,286 | 830,286 |
| Snow data for 20 stations in Russia | 622,325 | 622,325 | 622,325 |

Relatório do Observatório Campos Rodrigues—Mozambique, 1909–60. Two sets of publications were found for Mozambique at the Instituto Dom Luiz library. The first collection, the one with the oldest data, contains the observations for Maputo (then Lourenço Marques, Campos Rodrigues Observatory) for 1909–60 (bihourly from 1910). The data were imaged, and the 1910–14 period and the 1947–60 period (in addition to the 1915–46 period in *Anais das Colonias* recovered previously in ERA-CLIM) were digitized. This is a very complete set of data containing the same variables as the Angola set mentioned above, as well as soil temperatures (with depth reaching 1.5 m below ground level) and grass temperatures. Priority was given to surface/MSL pressure, and for the 1947–60 period all data (recorded every 2 h) have been digitized and sent to ISPD.

Província de Moçambique—Boletim Mensal, 1934–56. The other Mozambique dataset was found to contain subdaily data for Maputo only, for 1934–44, and for several stations in the 1951–56 period. There is a gap between 1945 and 1950 in the Instituto Dom Luiz collection of these publications. Besides imaging the whole set, it was chosen to digitize the 1951–56 period for the following stations: Beira, Inhambane, Quelimane, Mossuril, Murrebué, Tete, Vila Cabral, and Lumbo. The same set of observed variables was recovered as for the Angola set mentioned above.

Registo de Todas las Observaciones del Tiempo de Superficie (Fuerza Aerea de Chile), 1950–58. In South America ERA-CLIM2 was given access to hourly surface observations performed by Fuerza Aerea de Chile for 1950–58 and for 41 stations. All data were handwritten and imaged and contained the following variables, among others: dewpoint temperature, cloud cover, wind direction and speed, surface pressure, dry-bulb temperature, accumulated daily

precipitation, daily maximum and minimum temperature, and relative humidity. Four values per day were digitized manually. As it was found that the year 1956 had no pressure data in ISPD for Chile, priority has been set to digitize this year, apply QC, format, and send the data to ISPD, a process that is nearing its end.

South China Sea region, 1933–41 (Xanghai logbooks). Since ERA-CLIM, when South China Sea land surface pressure contained in the logbooks from Hong Kong was digitized and formatted for ISPD for the 1894–1932 period, a new set of data were found (Xanghai logbooks) for 1933–41. ERA-CLIM2 digitized land surface pressure data for 77 stations in the South China Sea and river ports from Russia to China, which was subsequently quality controlled and formatted and sent to ISPD. The data are also available online (<http://eraclim2.rd.ciencias.ulisboa.pt/>).

Russians stations prior to 1965. Data from Russian stations were extended backward prior to 1966 from the beginning of observation (four observations per day). A subset of 246 stations of the possible 518 was selected for this backward extension and data were prepared. The following data operations were done for 246 stations to obtain the extension back in time from 1965. The work comprised the assessment of

existing sources of data on computer media in any format and in hardcopies to select subset of stations acceptable for period extension; the reformatting of old odd-formatted data to fit the common format for later data; filling gaps in data for the period before 1966 by digitizing hardcopy material, transforming digitized data to a common format; adjustment of time and date; and QC of data.

UPPER-AIR DATA. A particular focus of the project was on historical upper-air data (Stickler et al. 2014; Ramello Pralungo et al. 2014). ERA-CLIM2 continued to digitize and QC the rich sources discovered during ERA-CLIM. These efforts led to an update of the Comprehensive Historical Upper-Air Network (CHUAN; Stickler et al. 2010). In total (ERA-CLIM and ERA-CLIM2), 1.3 million station days' worth of upper-air profiles were cataloged (Table 2), most of which were digitized and quality controlled. Since stations typically have two soundings per day, with several variables at many levels, 1.3 million station days correspond to tens of millions of individual data points. A map of all stations, with an indication of the number of station days, is given in Fig. 1. Noteworthy is the digitization of the data from the so-called international days from the 1920s. This was an effort coordinated by the International Aerological Commission. On specific

days (typically once per month), balloons were launched from many stations worldwide. Although assimilating such data into a reanalysis may not help much, it is a superb opportunity for comparison. A further effort concerned upper-air data from ships, again continuing from ERA-CLIM (Stickler et al. 2015). In the following we give a brief overview of the most important sources.

Russian upper-air data prior to 1960. Operational radiosonde data from 41 stations from Russia (Fig. 1) prior to 1960 were digitized, which is a backward extension of existing data holdings (such as the Integrated Global Radiosonde

TABLE 2. Upper-air observations (No. of station days) digitized and processed within ERA-CLIM and ERA-CLIM2 (CHUANv2.1).

| Source | Imaged | Digitized | Quality controlled |
|--|---------|-----------|--------------------|
| Russian radiosonde data for 41 stations, 1938–64 | 167,401 | 167,401 | 167,401 |
| Daily weather report, Germany, 1903–34 | 118,020 | 117,837 | 114,187 |
| India Meteorological Department, 1928–36 | 113,779 | 113,779 | 113,779 |
| Indian daily weather report, 1938–42 | 106,379 | 106,379 | 106,379 |
| Pakistan daily weather report, 1949–56 | 101,714 | 101,714 | 101,714 |
| Daily weather report, Cairo, Egypt, 1920–53 | 68,184 | 68,184 | 68,184 |
| Aerological observations, Netherlands, 1909–40 | 59,054 | 59,054 | 59,054 |
| International days, 1923–28 | 53,883 | 53,883 | 53,725 |
| Monthly bulletin, Portugal (Mozambique), 1938–56 | 27,402 | 23,292 | 23,292 |
| Annual bulletin, Finland, 1919–34 | 19,583 | 19,583 | 19,583 |
| Moving upper-air data, 1888–1947 | 9,000 | 9,000 | 9,000 |
| Météo-France database, 1948–58 | 0 | 33,818 | 33,818 |
| Comptes Rendus, 1923–57 | 216,608 | 86,476 | 67,201 |
| Other French sources, 1900–57 | 98,894 | 35,375 | 25,210 |
| Spanish data, 1912–61 | 14,519 | 14,519 | 14,519 |
| Portuguese annals, 1939–72 | 27,093 | 27,093 | 27,093 |
| Other sources, upper air | 77,632 | 75,746 | 75,762 |

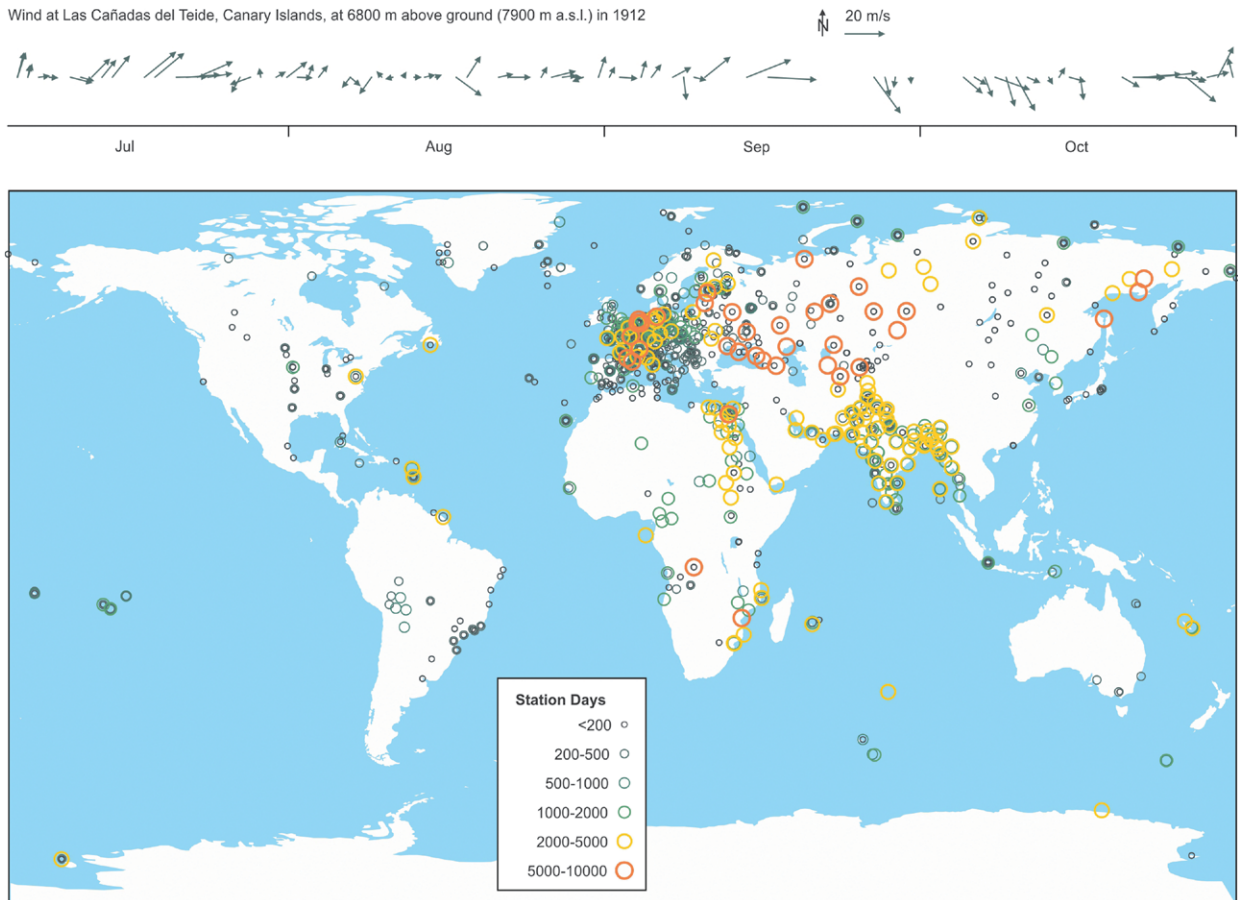


FIG. 1. (top) Wind at 7900 m above mean sea level obtained with pilot balloons at Las Cañadas del Teide on Tenerife, Canary Islands, in 1912. (bottom) Map of digitized upper-air data from ERA-CLIM and ERA-CLIM2. The circle size indicates the number of station days per station. In total, 1400 stations are shown.

Archive; see Durre et al. 2006) to the 1940s. The work comprised three parts:

- complementing data already existing in electronic format but which were incomplete (missing stations, missing levels, etc.);
- digitizing new sources from handwritten tables; and
- integrating computerized “views” of old punch cards in exotic formats (the handwritten tables were manually keyed by punching machines to paper punch cards in the 1970s and 1980s, and in the late 1980s the paper punch cards were copied to nine-track magnetic tapes “as is”). The effort in ERA-CLIM2 related to this source of data included decoding information from obsolete nine-track tape media and compiling a digital dataset.

The final dataset was compiled from these three sources. The data were formatted, quality controlled, and delivered to CHUAN, version 2.1. They were also

used for various analyses (Lavrov and Sterin 2017; Sterin and Lavrov 2017).

Central and northern European data, 1900s–30s and polar expeditions. The digitization of several large collections of upper-air data begun in ERA-CLIM was continued in ERA-CLIM2. An important source for early data was the German “daily weather report,” (1903–34), which contains data from many stations in the Atlantic–European sector as well as Russia. Also data from the Netherlands and Finland were digitized (Fig. 1). These three data sources provide rich information on the early period of aerological measurements (i.e., from the 1900s to the 1930s); in total, 200,000 station days were digitized. Also noteworthy are upper-air data from early Arctic expeditions such as Greenland expeditions in 1912/13 (Switzerland), 1926–31 (University of Michigan), and 1930/31 (Germany) or British, American, and Australian Antarctic expeditions.

Asian and African sources, 1920s–50s. A large amount of data were digitized from the reports of the India

Meteorological Department (1928–42) and the Pakistan daily weather report (1949–56). A further important source was upper-air data from Egypt (1920–53). These sources clearly leave a large imprint on the station map (Fig. 1) reflecting a colonial world at that time (Brönnimann and Wintzer 2018). As an example of a further Asian source, Fig. 2 shows a kite ascent from Beijing, China, on 12 April 1934. It illustrates the steps from the raw (graphical) recording to an evaluated table, which then was digitized. Upper-air data for the area of China and Korea from the 1930s and 1940s were digitized from the Sino–Swedish Hedin expedition (1930/31), the Chinese Academy (1932–35), and the Japanese occupation during the Second World War (Fig. 2). Very high-reaching upper-air profiles were digitized from the German East Africa expedition in 1908 (Brönnimann and Stickler 2013) and from Batavia, Indonesia.

Spanish upper-air data, 1912–17 and 1930–61. A collection of Spanish pilot balloon data for Madrid, Barcelona, La Coruña, Las Cañadas del Teide, and Izaña between 1912 and 1917 was digitized. Figure 1 (top) shows a short excerpt of a wind record at 7900 m above mean sea level from Las Cañadas del Teide.

The station location is interesting for studying the subtropical jet, elongated upper-trough systems, or transport of Saharan dust over the Atlantic Ocean. The Ebro station contained a much longer series of pilot balloon observations spanning 1930–61. All data have been digitized, quality controlled, and supplied to CHUAN, version 2.1. The corresponding metadata were added to the project’s registry (<http://eraclim-global-registry.fc.ul.pt/>).

French upper-air data, 1919–57. Upper-air data were located, inventoried, imaged, and digitized for 52 French stations, comprising 31 France mainland stations and 21 France overseas stations (New Caledonia, French Polynesia, French West Indies, and French Guyana; Table 2 and Fig. 1).

Pilot balloon data for the period 1919–57 were digitized from handwritten reports. Major efforts have been undertaken to locate records of upper-air data from France and French ex-colonies in 82 French archives, to image all the located records (one million images generated within ERA-CLIM and ERA-CLIM2), to create new long-term series of upper-air wind, and to compile metadata on these historical observations. Seven long-term series of upper-level

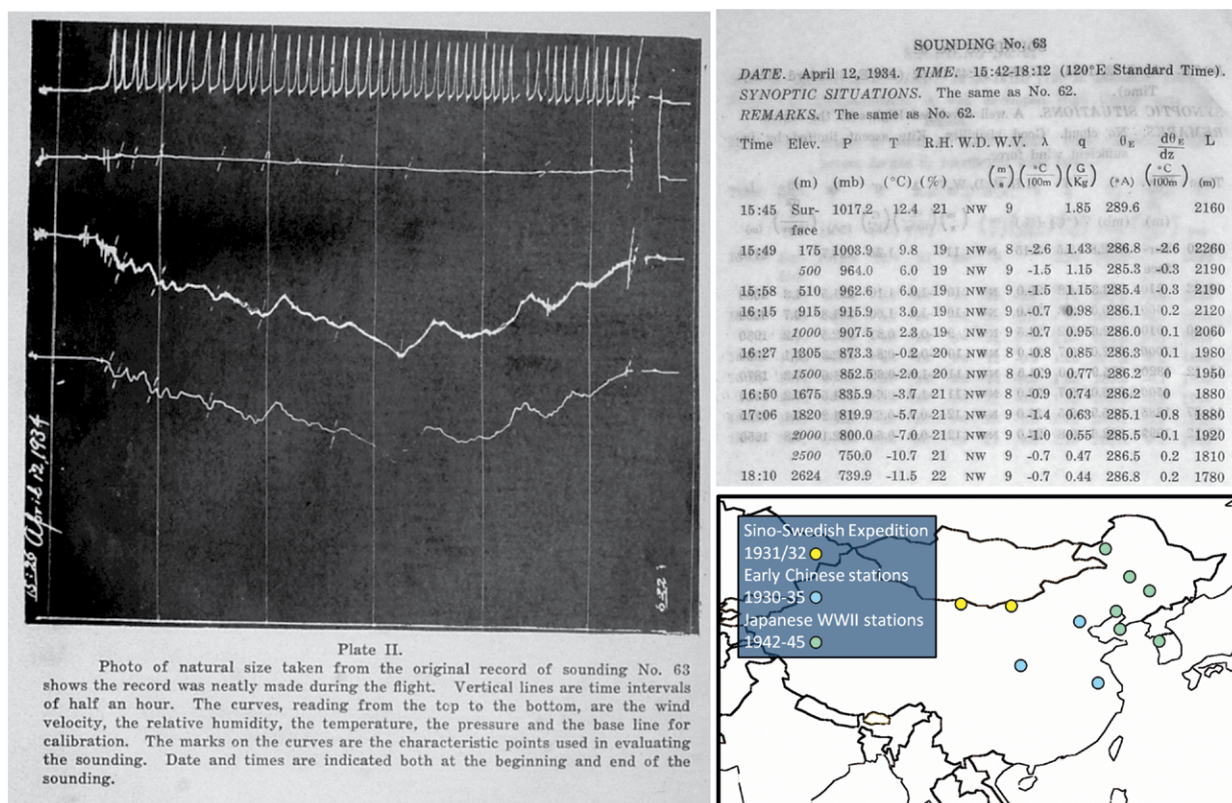


FIG. 2. Kite ascent from Beijing, 12 Apr 1934: (left) raw recording and (top right) tabulated data (from the *Bulletin of the Upper Air Current Observations*, Vol. I–VI, Academia Sinica, Nanking, China). (bottom right) The map shows the series from China and Korea from the 1930s and early 1940s digitized within ERA-CLIM and ERA-CLIM2.

wind for the French mainland and one series in Corsica longer than 25 years were generated: Ajaccio (1920–48), Antibes/Nice (1923–48), Bordeaux (1929–21, 1923–57), Le Bourget (1920–21, 1923–48), Lyon (1920–21, 1923–47), Perpignan (1920–21, 1923–57), Poitiers (1921, 1923–48), and Toulouse (1923–47). In addition, early radiosonde data for Trappes (Fig. 3) and Bordeaux for the period 1937–39 have been located in an old internal publication and have been digitized to extend the series provided within ERA-CLIM.

OCEANIC DATA. *Historical Southern Ocean data.* Our knowledge of climate variability and change in the Southern Ocean is currently limited by a lack of historical observations from that region. This limits the skill of reanalyses here because not only are there relatively few observations to assimilate but the boundary conditions are also relatively poorly understood. The Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST) analysis (Rayner et al. 2003) was further developed under ERA-CLIM and ERA-CLIM2, revising both the SST and sea ice components to version 2. Titchner and Rayner (2014) document the new sea ice analysis (which is available at www.metoffice.gov.uk/hadobs/hadisst2/). The representation of Antarctic sea ice remains fairly crude in this analysis prior to 1973, so a search was undertaken for more observations. This search was continued under ERA-CLIM2 and has yielded information not only on sea ice but also (primarily) on surface meteorological variables for the Southern Ocean for the late nineteenth and early twentieth centuries.

These data, mostly from ships' logbooks in the Southern Ocean, have been imaged but not keyed, from the original documents. These documents include dedicated meteorological logbooks as well as common ships' logbooks, meteorological forms, whale-catch books, day reports, ice reports, ice charts, and other relevant items. Types and frequency of observations will vary from one document to another, but typically there is barometric pressure and air temperature, wind direction and force, weather, cloud cover, SST, sea state and swell, sometimes salinity and biological observations, sea ice and icebergs, and other phenomena such as aurora. The logbooks and other documents imaged are mostly from the twentieth century up to 1950, with a few from the nineteenth century such as the Ross Antarctic expedition of 1839–43.

Observations have been imaged from naval vessels (British and Chilean), whaling vessels (British and Norwegian), vessels of exploration, research vessels, and merchant shipping (British and Finnish). In all about 131,000 images have been captured (58,000

from the United Kingdom, 30,500 from Norway, 20,300 from Finland, and 22,000 from Chile). The images are held at the Met Office. This is expected to provide more than one million observations, once the data are keyed. Much of the British and Chilean data are for the Weddell Sea area and around the Antarctic Peninsula. The Norwegian whaling data cover the Weddell Sea and Indian Ocean sectors of the Antarctic. The data from merchant shipping (1900–39, 1946–50) include high-latitude transits of the South Pacific from Australia and New Zealand toward Cape Horn or the west coast of South America by British and Finnish vessels. The collection of 131,000 images does not represent the entirety of historic data available for the Southern Ocean. There are many more known sources of ships' logbooks in Norway and Finland. Archives in Sweden, Denmark, Germany, Argentina, Australia, New Zealand, Japan, and the United States have yet to be explored and documented.

Development of HadIOD. The Met Office Hadley Centre Integrated Ocean Database (HadIOD; Atkinson et al. 2014) was first developed under ERA-CLIM and further improved under ERA-CLIM2 to produce version 1.2.0.0. HadIOD is a database of global ocean temperature and salinity observations, combining data from surface-only and profiling instruments, with quality flags, bias adjustments, and uncertainty estimates for each observation (where possible). Chief data sources are the International Comprehensive Ocean–Atmosphere Data Set (ICOADS) (Woodruff et al. 2011) for SST observations and the Met Office Hadley Centre EN series dataset (Good et al. 2013) for temperature and salinity profile observations. The ultimate aim of HadIOD is to provide input for coupled data assimilation.

HadIOD.1.2.0.0 has undergone the following recent developments: (i) greater time coverage (now starting in 1850); (ii) new data sources, for example, high-temporal-resolution global tropical moored buoy array data; (iii) improvements to bias corrections; (iv) a new output format for assimilation; and (v) improved quality control of data from ships and drifting buoys over their measuring lifetimes.

SNOW DATA. An effort to establish a global dataset of distributed snow observations resulted in a new collection of snow water equivalent (SWE) values measured at snow courses of the Northern Hemisphere. SWE recordings from distributed snow courses were obtained covering the regions of Russia (and the former Soviet Union), Canada,



| VALEURS DES ÉLÉMENTS POUR LES GÉOPOTENTIELS PRINCIPAUX (P T U/Φ; d/Φ). | | | | | | | | | | | |
|--|-----------------|--------------|--------------|-------------------|-------------------|--------------|--------------|-----------------|-------------------------------------|--|--|
| Φ | P | T | U | θ | γ | u | v | ψ | DESCRIPTION DU TEMPS | | |
| Soi | $P_s = 1013.25$ | $T_s = 15.0$ | $U_s = 10.0$ | $\theta_s = 30.0$ | $\gamma_s = 10.0$ | $u_s = 10.0$ | $v_s = 10.0$ | $\psi_s = 10.0$ | 1. Encre | | |
| 500 | 955 | -6.0 | 83 | | | | | | 2. Au moment du lancer (instant 0) | | |
| 1.000 | 897 | -6.0 | 85 | | | | | | 3. Au moment du lancer (instant 0) | | |
| 1.500 | 843 | -2.2 | 87 | | | | | | 4. Au moment du lancer (instant 0) | | |
| 2.000 | 790 | 0.0 | 90 | | | | | | 5. Au moment du lancer (instant 0) | | |
| 3.000 | 695 | -6.0 | 95 | | | | | | 6. Au moment du lancer (instant 0) | | |
| 4.000 | 610 | -12.2 | 100 | | | | | | 7. Au moment du lancer (instant 0) | | |
| 5.000 | 532 | -19.0 | 100 | | | | | | 8. Au moment du lancer (instant 0) | | |
| 6.000 | 464 | -25.2 | 100 | | | | | | 9. Au moment du lancer (instant 0) | | |
| 7.000 | 402 | -33.5 | | | | | | | 10. Au moment du lancer (instant 0) | | |
| 8.000 | 345 | -41.0 | | | | | | | 11. Au moment du lancer (instant 0) | | |
| 9.000 | 297 | -50.0 | | | | | | | 12. Au moment du lancer (instant 0) | | |
| 10.000 | 254 | -59.0 | | | | | | | 13. Au moment du lancer (instant 0) | | |
| 11.000 | 216 | -58.2 | | | | | | | 14. Au moment du lancer (instant 0) | | |
| 12.000 | | | | | | | | | 15. Au moment du lancer (instant 0) | | |
| 13.000 | | | | | | | | | 16. Au moment du lancer (instant 0) | | |
| 14.000 | | | | | | | | | 17. Au moment du lancer (instant 0) | | |
| 15.000 | | | | | | | | | 18. Au moment du lancer (instant 0) | | |

FIG. 3. (top) Balloon ascent in Trappes, France, around 1930. (bottom) Data sheet of radiosonde ascent on 21 Dec 1945.

and Finland. Thus, the new collection covers all major regions of seasonal snow cover located in the Northern Hemisphere. The dataset is based on the in situ observations by the Russia Research Institute

of Hydrometeorological Information/World Data Center, Environment Canada, and the Finnish Environment Institute. The length of snow courses varies from several kilometers (Finnish snow courses and

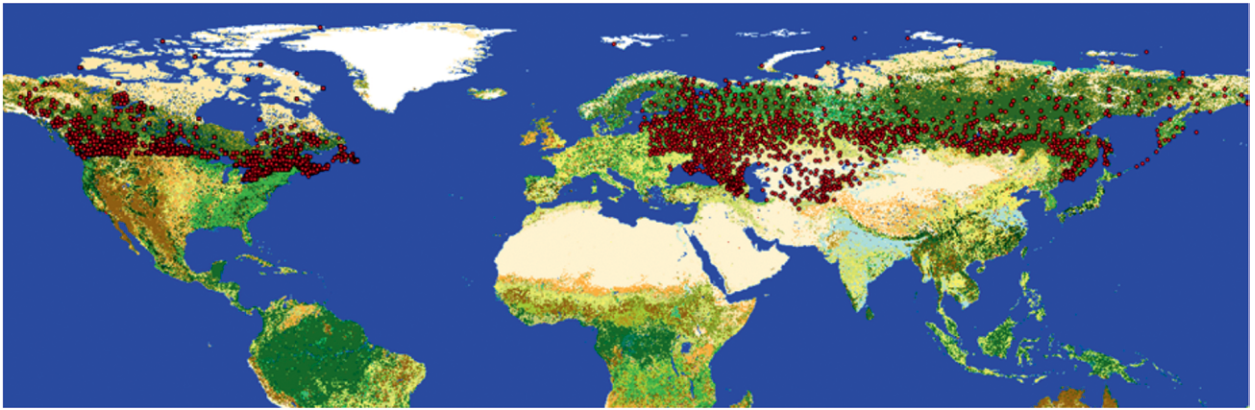


FIG. 4. Distributed snow course observation dataset of SWE collected in ERA-CLIM2.

open forest canopy/steppe courses in Russia) to a couple of hundred meters (Canadian snow courses and those of closed forests in Russia). The dataset extends from 1935 to 2014 at 3,589 locations (Fig. 4), the total number of observations being around 1.2 million. The SWE values were calculated from snow depth and density samples. Depth measurements were performed at up to 100 locations obtained at a constant interval through the snow course. The snowpack bulk density is sampled more sparsely, typically every tenth depth sample. In addition to SWE, the dataset has information on snow course center coordinates, mean snow depth, mean density, and the day of an individual observation. The typical frequency of observation at a single site varies from once a week to once a month, and the observations were carried out throughout the period of seasonal snow cover. Figure 4 shows the locations of included snow courses.

In addition to SWE, historical snow depth data from Russia were digitized and processed. Again, the digitization mainly extended the list of stations in the snow parameter dataset. Furthermore, 20 additional stations were digitized that were not available electronically. Owing to numerous metadata records on historical changes in snow observational practices in the former Soviet Union and on the changes in QC processing and quality flag codes, it became possible to merge the data into a consistent database. (The data are available from www.meteo.ru.)

Figure 5 shows climatologies of snow depth for the months of April, October, and November (Wegmann et al. 2017). The bottom row shows corresponding maps from the ECMWF twentieth-century reanalysis (ERA-20C; Poli et al. 2016), which does not assimilate snow depth data. The figure shows that ERA-20C realistically reproduces the broad features of climatological average snow depth.

HISTORICAL SATELLITE DATA. Satellite data are used in almost all global reanalyses covering the last 50–60 years. However, they only became available in the mid-1960s, and the satellites and instruments were mostly built for the purpose of weather monitoring. Their use in a reanalysis requires a quantification and correction of long-term trends due to changes in the characteristics of satellites and performance of sensors during their operational lifetime in space, as well as systematic differences between instruments of the same kind (e.g., in their spectral response). A reprocessing of the data that applies corrections for such effects is fundamental to serve the generation of physically consistent data records of geophysical variables by reanalysis. Consequently, in ERA-CLIM2, further improvement of reanalyses of the satellite era was contributed to by producing reprocessed and recalibrated satellite data records and by including more observations from early satellite instruments. An assessment of satellite sounder radiances suitable for use in reanalyses, mainly from the 1970s, was made for ERA-CLIM2 and is documented in Table 3.

Specific efforts addressed four different satellite time series:

- Radiances from the Meteosat Visible and Infrared Imager (MVIRI) and the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instruments aboard the Meteosat first- and second-generation satellites were recalibrated, that is, referenced to more accurate measurements from instruments in polar orbit, in particular from the Infrared Atmospheric Sounding Interferometer (IASI), Atmospheric Infrared Sounder (AIRS), and High Resolution Infrared Radiation Sounder (HIRS) instruments. The recalibrated Meteosat radiances were used to derive clear-sky and all-sky

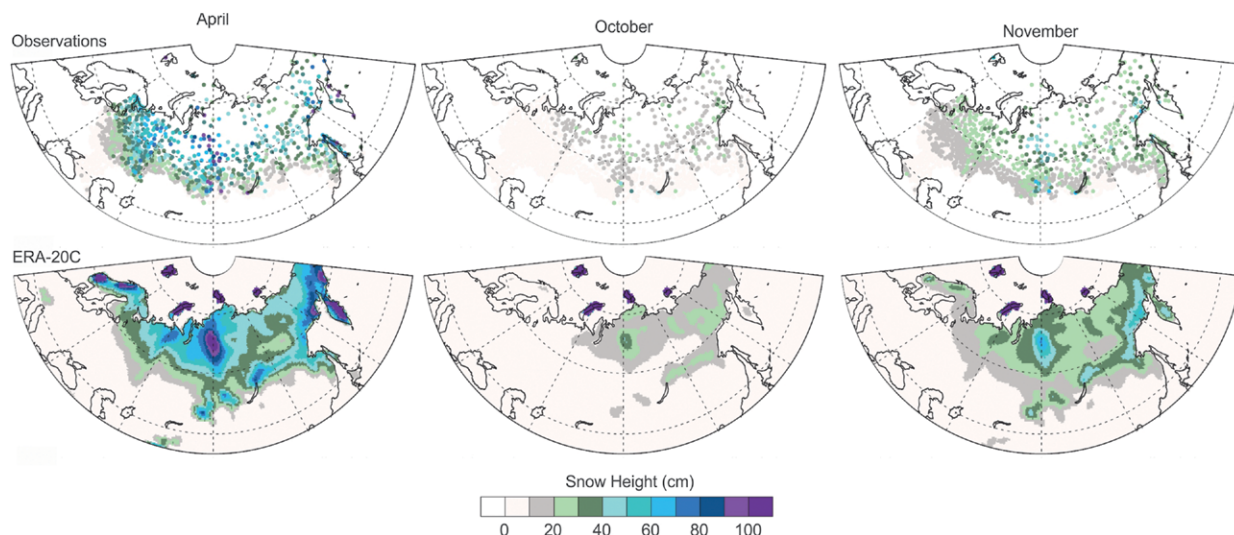


FIG. 5. Snow depth climatology for (left) April, (center) October, and (right) November 1981–2010, from (top) station data from the former Soviet Union and (bottom) ERA-20C (from Wegmann et al. 2017).

radiances and atmospheric motion vectors (Borde et al. 2014) for the period 1982–2014. The recalibration of geostationary radiances is a complex process involving multiple steps: 1) collection of geostationary and reference instruments datasets; 2) selection of reference instruments and bringing the reference data to match spectrally with the geostationary measurements, which involves rigorous data analysis and research; 3) finding collocations between the geostationary and reference data (which is computationally expensive);

and 4) use of the collocated measurements to determine the recalibration coefficients. Figure 6 shows the cross-calibrated Meteosat infrared (IR) window channel ($10.5\text{--}12.5\text{ }\mu\text{m}$) radiances for a portion of the time series from 1991 to 2006. The time series shows an almost perfect alignment of the resultant radiances from different Meteosat satellites, and the difference plot demonstrates the successful elimination of some artifacts in the time series and the removal of an artificial trend from the data.

TABLE 3. Additional infrared and microwave sounders considered suitable for reanalyses.

| Sensor | Platform(s) | Period | Primary variables retrievable | Status |
|---|--|-----------|--|---|
| Pressure Modulator Radiometer (PMR) | <i>Nimbus-6</i> | 1975–76 | Stratospheric temperature | Data available at University of Oxford and Met Office, copy at ECMWF |
| HIRS | <i>Nimbus-6</i> | 1975–76 | Temperature and humidity profiles | Data available at NASA, copy at ECMWF |
| MVIRI | <i>Meteosat-1</i> – <i>Meteosat-4</i> | 1978–95 | Surface and cloud-top temperature and upper-tropospheric humidity | Data available at EUMETSAT |
| Infrared Interferometer Spectrometer (IRIS) | <i>Nimbus-3</i> and <i>Nimbus-4</i> | 1970–71 | Temperature and humidity profiles and trace gases | Data available at NASA for <i>Nimbus-4</i> , copy at ECMWF |
| Vertical Temperature Profile Radiometer (VTPR) | <i>NOAA-2</i> – <i>NOAA-5</i> | 1972–79 | Temperature and humidity profiles | Used already in ERA-40 and Japanese 55-year Reanalysis (JRA-55) |
| Scanning Multichannel Microwave Radiometer (SMMR) | <i>Nimbus-7</i> | 1975–83 | Sea ice, total column water vapor, ocean surface wind, and cloud liquid water path | New fundamental climate data record (FCDR) produced by EUMETSAT Satellite Application Facility on Climate Monitoring (CMSAF; www.cmsaf.eu) |
| SSM/T-2 | DMSP F8– F15 | 1992–2008 | Upper-tropospheric humidity | FCDR available at EUMETSAT |

- The Advanced Very High Resolution Radiometer (AVHRR) infrared window channel instrument data at $11\ \mu\text{m}$ were used to derive atmospheric motion vectors over areas poleward of 60° on the Northern and Southern Hemispheres using an algorithm from Borde et al. (2016). The AVHRR time series was stitched together from NOAA-7 to NOAA-19 and European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) *Meteorological Operational Satellite-A (MetOp-A)/MetOp-B* AVHRR measurements covering 1982–2015.
- Passive microwave sounder water vapor measurements with three channels in the 183-GHz water vapor line from three different instruments were cross calibrated using the Observation Feedback Archive (OFA) created by ECMWF during the ERA-CLIM project. These instruments were the following: (i) the Special Sensor Microwave/Temperature 2 (SSM/T2) instruments flown on four Defense Meteorological Satellite Program (DMSP) satellites (*F11*, *F12*, *F14*, and *F15*) covering 1992–2008; (ii) the Advanced Microwave Sounding Unit B (AMSU-B) aboard NOAA-15 to NOAA-17 covering 1999–2010; and (iii) the Microwave Humidity Sounder (MHS) aboard NOAA-18 and NOAA-19 and EUMETSAT *MetOp-A/MetOp-B* covering 2005–15. The OFA provides monitoring results for each of the instruments against the ECMWF interim reanalysis (ERA-Interim) and allows a relative cross calibration of all sounding instruments to one reference instrument in the time series [for SSM/T2 see Kobayashi et al. (2017)].
- Radio occultation (RO) bending-angle profiles were derived from the EUMETSAT *MetOp-A/MetOp-B* combined global positioning system (GPS)–global navigation satellite system (GLONASS) (GNSS) Receiver for Atmospheric Sounding (GRAS)

covering 2006–15; the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) covering 2006–15; and the Challenging Minisatellite Payload (CHAMP) covering 2001–10 missions. RO measurements are very accurate and help to constrain atmospheric temperature in the model throughout the atmosphere. The reprocessing of the GRAS data has resulted in approximately three million globally distributed bending-angle profiles.

The scope of effort needed for satellite data rescue and reprocessing is no less daunting than that for in situ data rescue (Poli et al. 2017). As part of satellite data rescue activities, advice on the use of historical satellite datasets was provided. Even more important, the capability to simulate the measurements of historical satellite instruments to allow monitoring and assimilation of radiances from such instruments has been enhanced. The Radiative Transfer for the Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) model (RTTOV; Saunders et al. 1999) has been updated since ERA-Interim was produced to include more of the historical sensors. In particular biases of individual

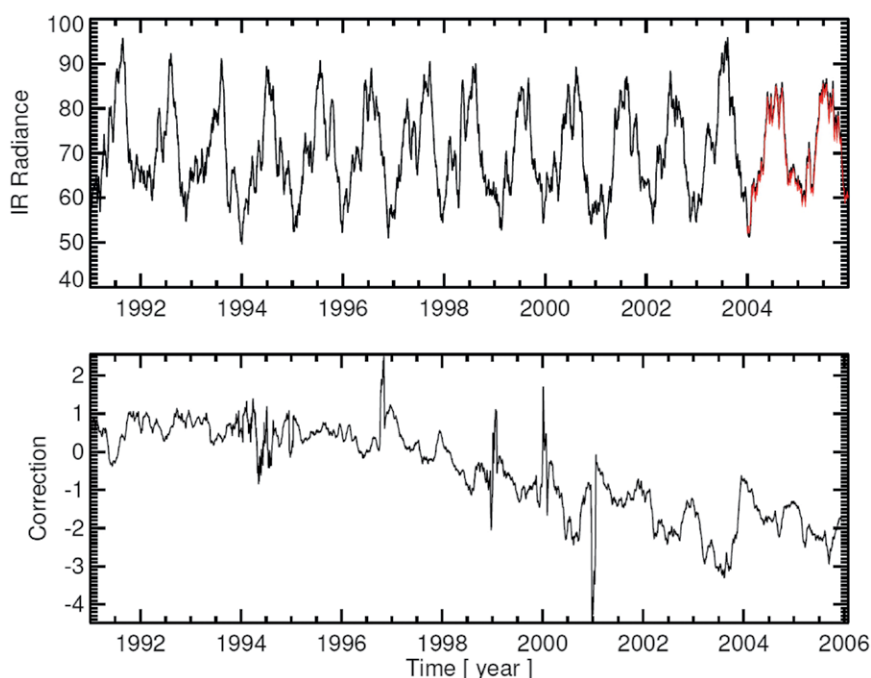


FIG. 6. (top) Recalibrated Meteosat IR channel radiances ($\text{mW m}^{-2} \text{sr}^{-1} \text{cm}^{-1}$) over Payerne, Switzerland, and (bottom) the difference between corrected radiances and those computed using operational calibration parameters. The black line shows Meteosat first-generation MVIRI radiances, and the red line shows Meteosat second-generation SEVIRI radiances. The cross-calibrated radiances have been homogenized using the *Meteosat-5* MVIRI spectral response function.

instruments were explored, and the work has shown that some spectral responses previously assumed in the radiative transfer model needed updating. The new RTTOV-12 model reduces some of the biases due to instrumental effects and also allows more accurate calculations with an increased number of pressure levels through the atmosphere. For example, the cell pressure of the Stratospheric Sounding Unit (SSU) reduced after launch and new coefficients have been developed to account for this (Nash and Saunders 2015). Coefficients for early satellite instruments have also been updated assuming the atmospheric CO₂ concentration for an earlier period, rather than using current CO₂ concentrations.

DATA MANAGEMENT, DATA USE, AND DATA PRODUCTS.

For the digitizing process and management of surface, ship, and upper-air data, a metadata database was developed within ERA-CLIM and ERA-CLIM2. This database is public (<http://eraclim-global-registry.fc.ul.pt/>) and will be helpful also to other projects as it contains metadata on thousands of surface and upper-air stations as well as ship logs. In addition to the station names, various station numbers, coordinates, and time period of measurements, the database contains information on the archive, variables covered, the state of the process, or data level (imaged, digitized, formatted, quality controlled, etc.) and offers the possibility to directly link to the data. This registry will allow others to make full use of the information compiled within ERA-CLIM and ERA-CLIM2 and to share their own data rescue information with the community.

In addition to updating existing repositories (see next section), observations rescued and processed in ERA-CLIM and ERA-CLIM2 were also directly used for the generation of reanalyses. Surface pressure observations were assimilated into both ERA-20C and the Coupled ECMWF Reanalysis of the Twentieth

Century (CERA-20C; Laloyaux et al. 2018), and some of the processed satellite data were assimilated into the fifth major global reanalysis produced by ECMWF (ERA5; Hersbach and Dee 2016). The satellite data are and will become available from EUMETSAT (under <http://navigator.eumetsat.int>).

As an example, Fig. 7 shows the case of a typhoon on 18 September 1906 that hit Hong Kong. ERA-20C, which only included one station within the region displayed (Hong Kong), did not reproduce the typhoon. Within ERA-CLIM, data from the South China Sea were digitized (large dots), and they were assimilated into CERA-20C. The CERA-20C ensemble mean produces a low pressure system, albeit too weak. However, ensemble member 9 of CERA-20C shows a stronger cyclone. In ERA-CLIM2, the South China Sea dataset has now been extended substantially, so that future reanalyses will be able to better simulate typhoons over South Asia in the early years.

Most of the observations digitized and processed within ERA-CLIM and ERA-CLIM2 were not assimilated into ERA-20C or CERA-20C, but they were used for their evaluation (e.g., Stickler et al. 2015). In addition, a test reanalysis was performed [European Reanalysis of the presatellite era (ERA-PreSAT); Hersbach et al. 2017], covering the period 1939–67 and assimilating the CHUAN, version-1, upper-air data (Stickler et al. 2010), demonstrating the potential of using historical upper-air data for reanalyses. Ongoing and future reanalysis efforts will be able to make use of the millions of profiles digitized as well as of the experience in generating ERA-PreSAT. Likewise, snow observations are not directly assimilated but were used to evaluate the reanalyses and their respective land surface products (Fig. 5).

CONTRIBUTIONS TO PUBLIC REPOSITORIES. The ERA-CLIM2 observations feed into existing repositories and are already used in other projects.

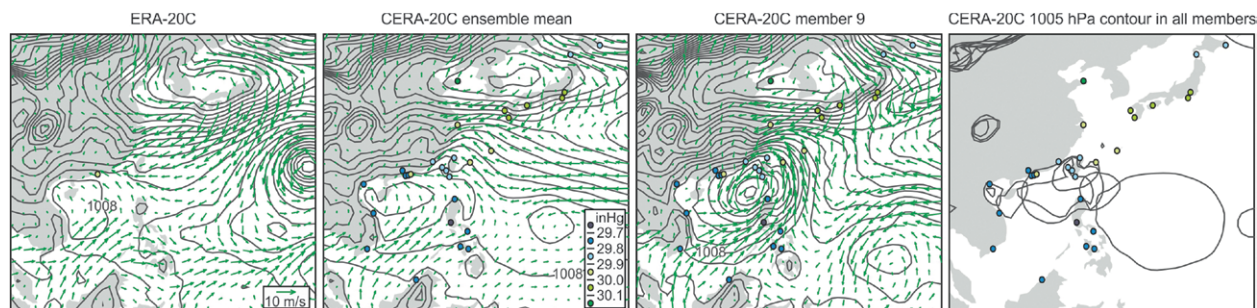


FIG. 7. Sea level pressure and 10-m wind during a typhoon at 1800 UTC 18 Sep 1906 from (left) ERA-20C and (from left center to right) CERA-20C ensemble mean, member 9, and map of the 1005-hPa contour for all members. The dots indicate pressure observations from the South China Sea dataset digitized by ERA-CLIM, which did not feed into ERA-20C but were assimilated into CERA-20C.

This section gives a brief overview (Table 4 lists the contributions to existing repositories). By submitting data to existing repositories (from where they are retrieved for reanalyses projects), they become part of their update cycles. This guarantees that version numbers are unique and that older versions remain available. When reanalyses approach real time, observations will be taken from the real-time data stream and not from these repositories, but the data will still be available to the user via the OFA.

The digitized surface data were submitted to ISPD (pressure) and ISTI (temperature). The ERA-CLIM2 upper-air data have been integrated into the latest version 2.1 of the CHUAN [database [for version 2 see Stickler et al. (2014)], which is ASCII formatted (and can be downloaded from <ftp://giub-torrent.unibe.ch/eraclim2/>). Those data have been reformatted into the ECMWF Observation Database (ODB2) format and have been stored in the Meteorological Archival and Retrieval System (MARS) archive. In this format the data can be immediately read by the Integrated Forecast System.

CHUAN2.1 will be used in the back extension of the Copernicus Climate Change Services (C3S) ERA5 (<http://apps.ecmwf.int/datasets/>) from 1950, where it will provide an essential source of upper-air information. After assimilation the upper-air data

will also be complemented with analysis feedback data such as background departure statistics and will be conveniently accessible via the Copernicus Data Store (<https://climate.copernicus.eu/climate-data-store>). Bias corrections at least for upper-air temperature back to 1950 have been developed as well and are described in detail in L. Haimberger et al. (2017, unpublished manuscript).

These procedures largely accomplish the requirements stated in a workshop that was held in Reading, United Kingdom (29 June–1 July 2015; www.ecmwf.int/sites/default/files/COP-CO-WS-Summary.pdf), on climate observation requirements. This includes the so-called observation feedback, which contains the original observations along with data assimilation quality flags, potential bias corrections, and departures from the reanalysis record. This represents a mine of information of interest to observation experts and reanalysis producers. Observation feedback is available from a number of reanalyses at present [Twentieth Century Reanalysis (20CR) and ERA-20C] and ensures that the full observation input is available.

MAKING THE EFFORT SUSTAINABLE. An important aspect of the work on observations is its sustainability. Future reanalysis projects require not

TABLE 4. ERA-CLIM and ERA-CLIM2 contributions to public repositories.

| Dataset/repository | Data contribution | URL | Reference (version) |
|--|---|---|--|
| MARS | Upper-air and surface meteorological data | Access through ECMWF | — |
| Hadley Centre Observation Datasets (HadOBS) | Sea surface temperature and sea ice in HadISST.2.2.0.0 | Sea ice available from www.metoffice.gov.uk/hadobs/ | Titchner and Rayner (2014) (HadISST.2.2.0.0 sea ice) |
| HadOBS | Ocean surface and profiles in HadIOD.1.2.0.0 | Available on request from www.metoffice.gov.uk/hadobs/ | Atkinson et al. (2014) (HadIOD.1.0.0.0) |
| HadOBS | Subdaily meteorological data (surface, all variables) in HadISD.2.0.0 | www.metoffice.gov.uk/hadobs/ | Dunn et al. (2016) (HadISD.2.0.0) |
| ISTI | Surface temperature | www.surface temperatures.org/ | Rennie et al. (2014) (ISTI, version 1) |
| ISPD | Pressure | https://reanalyses.org/index.php/observations/international-surface-pressure-databank | Cram et al. (2015) (ISPD, version 3) |
| CHUAN2.1 | Upper air | http://giub-torrent.unibe.ch/eraclim2 | Stickler et al. (2014) (CHUAN2.0) |
| Russian snow data | Snow cover and depth | www.meteo.ru | In preparation |
| Long-term SWE data record | SWE | http://litdb.fmi.fi/eraclim2.php | In preparation |
| European Union Surface Temperature for All Corners of Earth (EUSTACE) global station dataset | Surface temperature | Forthcoming | In preparation |

only observations but also metadatabases and data holdings (Thorne et al. 2017), data rescue services (Brönnimann et al. 2018, manuscript submitted to *Geosci. Data J.*), and an active climate data community. The C3S contributes to make this undertaking sustainable via investment in services on data rescue and global database production.

C3S is one of several dedicated information services recently established by the European Commission as part of its Copernicus Earth Observation Programme. The purpose of C3S is to provide open and free access to data, tools, and information needed to support climate adaptation and mitigation in Europe. Improving access to climate observations and reanalyses is a key objective. C3S is being implemented as an operational service similar to weather forecasting, which means that all data and products will be fully supported, quality assured, and reliably delivered on a well-defined schedule. This provides a unique opportunity to transfer some of the essential work on observations and reanalysis undertaken in ERA-CLIM and ERA-CLIM2 and related projects to a sustainable operational environment.

Various technical aspects of the work on observations described in this article are now included in C3S. For example, a C3S data rescue service is being developed that will build on the work started in ERA-CLIM on the collection and sharing of metadata on data rescue projects (Brönnimann et al. 2018, manuscript submitted to *Geosci. Data J.*). This service will also develop new tools for data rescue and promote best practices and in some cases provide direct support to data rescue communities (e.g., in the form of capacity-building workshops). C3S is also taking steps to improve access to observations in existing climate data archives, by supporting the development and maintenance of a new global, quality-controlled database of historic surface observations, including land and marine data, building on existing holdings at NOAA/NCEI and many other sources (Thorne et al. 2017). Similar activities will be initiated to ensure the maintenance and further development of a merged database of available upper-air meteorological observations, including all those collected in the ERA-CLIM and ERA-CLIM2 projects. EUMETSAT is continuing its recalibration and reprocessing activities within a dedicated C3S project in support of the next reanalysis of the satellite era.

Production of global and regional model-based reanalyses, which until recently has relied almost exclusively on research funding, is now conducted as a fully supported operational activity under Copernicus. This includes ERA5, the latest ECMWF

reanalysis, which will be completed in 2018, and incorporates various reprocessed satellite data records developed in ERA-CLIM. The next generation, ERA6, will likely be a coupled atmosphere–ocean reanalysis using data and methods developed in ERA-CLIM2.

Clearly, establishing reanalysis as an operational activity will bring many benefits to users and, as a result, will generate more support for work on data recovery and reprocessing. Nevertheless, a strong need remains for underpinning research on data and methodology to ensure that the state of the art continues to evolve.

CONCLUSIONS. Global dynamical reanalyses of the atmosphere and ocean rely on observations in various steps along the production chain. In addition to being assimilated into a reanalysis system, observations are used to constrain the model boundary conditions, for uncertainty determination of other observations, and for the evaluation of data products. In ERA-CLIM2, major efforts were devoted to observations, namely, data rescue, processing and management, compilation and quality control, and error estimation:

- New sea surface and sea ice datasets were produced that serve as boundary conditions for reanalyses.
- Hundreds of thousands of images were taken within ERA-CLIM and ERA-CLIM2 and 6.4 million station days were digitized, amounting to tens of millions of data points.
- Several snow products were generated, encompassing SWE and snow depth.
- Satellite data were recalibrated and reprocessed to provide consistent satellite data records from the 1980s until today.
- The HadIOD and HadISD databases of quality-controlled surface and ocean measurements were extended and improved.

A large fraction of the data were not assimilated into the reanalysis yet but fed into global repositories on which future reanalysis efforts will be based. The work on observations therefore is one cycle ahead of reanalysis production. The continuation of this effort is vital for the success of future reanalysis projects.

ACKNOWLEDGMENTS. ERA-CLIM2 was funded by EU-FP7 (Grant Agreement 607029). The project collaborated closely with the Horizon2020 project EUSTACE (640171), as well as the Swiss National Science Foundation projects DECADE (147320) and CHIMES (169676).

REFERENCES

- Allan, R., P. Brohan, G. P. Compo, R. Stone, J. Luterbacher, and S. Brönnimann, 2011: The international Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. *Bull. Amer. Meteor. Soc.*, **92**, 1421–1425, <https://doi.org/10.1175/2011BAMS3218.1>.
- Atkinson, C. P., N. A. Rayner, J. J. Kennedy, and S. A. Good, 2014: An integrated database of ocean temperature and salinity observations. *J. Geophys. Res. Oceans*, **119**, 7139–7163, <https://doi.org/10.1002/2014JC010053>.
- Blišňák, V., M. A. Valente, and J. Bethke, 2015: Homogenization of time series from Portugal and its former colonies for the period from the late 19th to the early 21st century. *Int. J. Climatol.*, **35**, 2400–2418, <https://doi.org/10.1002/joc.4151>.
- Borde, R., M. Doutriaux-Boucher, G. Dew, and M. Carranza, 2014: A direct link between feature tracking and height assignment of operational EUMETSAT atmospheric motion vectors. *J. Atmos. Oceanic Technol.*, **31**, 33–46, <https://doi.org/10.1175/JTECH-D-13-00126.1>.
- , O. Hautecoeur, and M. Carranza, 2016: EUMETSAT global AVHRR winds product. *J. Atmos. Oceanic Technol.*, **33**, 429–438, <https://doi.org/10.1175/JTECH-D-15-0155.1>.
- Brönnimann, S., and A. Stickler, 2013: Aerological observations in the tropics in the early twentieth century. *Meteor. Z.*, **22**, 349–358, <https://doi.org/10.1127/0941-2948/2013/0458>.
- , and J. Wintzer, 2018: Use imprint of society and history on climate data to inform climate services. *Nature*, **554**, 423, <https://doi.org/10.1038/d41586-018-02201-z>.
- Buizza, R., and Coauthors, 2018: The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of Earth-system climate reanalyses. *Bull. Amer. Meteor. Soc.*, **99**, 1003–1014, <https://doi.org/10.1175/BAMS-D-17-0199.1>.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, <https://doi.org/10.1002/qj.776>.
- Cram, T. A., and Coauthors, 2015: The International Surface Pressure Databank version 2. *Geosci. Data J.*, **2**, 31–46, <https://doi.org/10.1002/gdj3.25>.
- Dunn, R. J. H., K. M. Willett, D. E. Parker, and L. Mitchell, 2016: Expanding HadISD: Quality controlled, sub-daily station data from 1931. *Geosci. Instrum. Methods Data Syst.*, **5**, 473–491, <https://doi.org/10.5194/gi-5-473-2016>.
- Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the integrated global radiosonde archive. *J. Climate*, **19**, 53–68, <https://doi.org/10.1175/JCLI3594.1>.
- Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans*, **118**, 6704–6716, <https://doi.org/10.1002/2013JC009067>.
- Hersbach, H., and D. Dee, 2016: ERA5 reanalysis is in production. *ECMWF Newsletter*, No. 147, ECMWF, Reading, United Kingdom, 5–6, www.ecmwf.int/sites/default/files/elibrary/2016/16299-newsletter-no147-spring-2016.pdf.
- , and Coauthors, 2017: The potential value of early (1939–1967) upper-air data in atmospheric climate reanalysis. *Quart. J. Roy. Meteor. Soc.*, **143**, 1197–1210, <https://doi.org/10.1002/qj.3040>.
- Hunziker, S., and Coauthors, 2017: Identifying, attributing, and overcoming common data quality issues of manned station observations. *Int. J. Climatol.*, **37**, 4131–4145, <https://doi.org/10.1002/joc.5037>.
- Kobayashi, S., P. Poli, and V. O. John, 2017: Characterisation of Special Sensor Microwave Water Vapor Profiler (SSM/T-2) radiances using radiative transfer simulations from global atmospheric reanalyses. *Adv. Space Res.*, **59**, 917–935, <https://doi.org/10.1016/j.asr.2016.11.017>.
- Laloyaux, P., and Coauthors, 2018: CERA-20C: A coupled reanalysis of the Twentieth Century. *J. Adv. Model. Earth Syst.*, **10**, 1172–1195, <https://doi.org/10.1029/2018MS001273>.
- Lavrov, A., and A. Sterin, 2017: Comparison of free atmosphere temperature series from radiosonde and satellite data. *Russ. Meteor. Hydrol.*, **42**, 95–104, <https://doi.org/10.3103/S1068373917020030>.
- Lawrence, H., and Coauthors, 2017: An evaluation of FY-3C MWRI and assessment of the long-term quality of FY-3C MWHS-2 at ECMWF and the Met Office. ECMWF Tech. Memo. 798, 26 pp.
- Nash, J., and R. Saunders, 2015: A review of Stratospheric Sounding Unit radiance observations for climate trends and reanalyses. *Quart. J. Roy. Meteor. Soc.*, **141**, 2103–2113, <https://doi.org/10.1002/qj.2505>.
- Poli, P., and Coauthors, 2016: ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Climate*, **29**, 4083–4097, <https://doi.org/10.1175/JCLI-D-15-0556.1>.
- , and Coauthors, 2017: Recent advances in satellite data rescue. *Bull. Amer. Meteor. Soc.*, **98**, 1471–1484, <https://doi.org/10.1175/BAMS-D-15-00194.1>.
- Ramella Pralungo, L., L. Haimberger, A. Stickler, and S. Brönnimann, 2014: A global radiosonde and tracked balloon archive on 16 pressure levels (GRASP) back to 1905—Part 1: Merging and interpolation to 00:00 and 12:00 GMT. *Earth Syst. Sci. Data*, **6**, 185–200, <https://doi.org/10.5194/essd-6-185-2014>.

- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, <https://doi.org/10.1029/2002JD002670>.
- Rennie, J. J., and Coauthors, 2014: The international surface temperature initiative global land surface databank: Monthly temperature data release description and methods. *Geosci. Data J.*, **1**, 75–102, <https://doi.org/10.1002/gdj3.8>.
- Saunders, R. W., M. Matricardi, and P. Brunel, 1999: An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, **125**, 1407–1425, <https://doi.org/10.1002/qj.1999.49712555615>.
- Smith, A., N. Lott, and R. Vose, 2011: The integrated surface database: Recent developments and partnerships. *Bull. Amer. Meteor. Soc.*, **92**, 704–708, <https://doi.org/10.1175/2011BAMS3015.1>.
- Sterin, A., and A. Lavrov, 2017: On the estimates of tropospheric temperature anomalies in 2015–2016. *Fundam. Appl. Climatol.*, **2**, 111–129, <https://doi.org/10.21513/2410-8758-2017-2-110-128>.
- Stickler, A., and Coauthors, 2010: The Comprehensive Historical Upper-Air Network. *Bull. Amer. Meteor. Soc.*, **91**, 741–751, <https://doi.org/10.1175/2009BAMS2852.1>.
- , and Coauthors, 2014: ERA-CLIM: Historical surface and upper-air data for future reanalyses. *Bull. Amer. Meteor. Soc.*, **95**, 1419–1430, <https://doi.org/10.1175/BAMS-D-13-00147.1>.
- , and Coauthors, 2015: Upper-air observations from the German Atlantic Expedition (1925–27) and comparison with the Twentieth Century and ERA-20C reanalyses. *Meteor. Z.*, **24**, 525–544, <https://doi.org/10.1127/metz/2015/0683>.
- Thorne, P., and Coauthors, 2017: Toward an integrated set of surface meteorological observations for climate science and applications. *Bull. Amer. Meteor. Soc.*, **98**, 2689–2702, <https://doi.org/10.1175/BAMS-D-16-0165.1>.
- Titchner, H. A., and N. A. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res. Atmos.*, **119**, 2864–2889, <https://doi.org/10.1002/2013JD020316>.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, <https://doi.org/10.1256/qj.04.176>.
- Wegmann, M., Y. Orsolini, E. Dutra, O. Bulygina, A. Sterin, and S. Brönnimann, 2017: Eurasian snow depth in long-term climate reanalyses. *Cryosphere*, **11**, 923–935, <https://doi.org/10.5194/tc-11-923-2017>.
- Woodruff, S. D., and Coauthors, 2011: ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967, <https://doi.org/10.1002/joc.2103>.